

Those skilled in the art should appreciate that they can readily use the disclosed conception and specific embodiments as a basis for designing or modifying other structures for carrying out the same purposes of the present invention without departing from the spirit and scope of the invention as defined by the appended claims.

We claim:

1

- 1 1. A method for buffering packets transmitted to an
2 Infiniband port by an Infiniband device linked to the
3 port, comprising:
4 providing a portion of a memory for buffering the
5 packets, wherein the portion has a size A;
6 transmitting flow control credits to advertise to the
7 device buffering resources of a size B, wherein B
8 is greater than A;
9 determining when the portion is filled with a
10 predetermined amount of the packets; and
11 transmitting flow control credits to the device to stop
12 transmission of the packets in response to said
13 determining.
- 1 2. The method of claim 1, wherein said transmitting flow
2 control credits to advertise to the device buffering
3 resources of a size B comprises transmitting flow
4 control credits to the device for a plurality of
5 Infiniband virtual lanes configured on the port.
- 1 3. The method of claim 2, wherein said plurality of
2 Infiniband virtual lanes comprises a number of data
3 virtual lanes from the list consisting of fifteen,
4 eight, four and two.

1 4. The method of claim 1, wherein said transmitting flow
2 control credits to the device to stop transmission of
3 the packets comprises transmitting flow control credits
4 to the device for a plurality of Infiniband virtual
5 lanes configured on the port.

1 5. The method of claim 1, further comprising:
2 providing a second memory for buffering the packets
3 transmitted subsequent to said determining.

1 6. The method of claim 5, wherein said second memory is
2 coupled between the port and the first memory.

1 7. The method of claim 5, wherein said determining the
2 portion is filled a predetermined amount comprises
3 determining the portion is approximately full.

1 8. The method of claim 5, wherein said providing a second
2 memory comprises providing a second memory having a
3 size C.

1 9. The method of claim 8, wherein said size C is based on an
2 amount of data that may be transmitted to the port
3 during a latency time required to stop transmission of
4 the packets in response to said determining.

1 10. The method of claim 9, wherein said latency time
2 comprises an approximate amount of time required to

3 perform said transmitting flow control credits to the
4 device to stop transmission of the packets in response
5 to said determining.

1 11. The method of claim 10, wherein said transmitting flow
2 control credits to the device to stop transmission of
3 the packets in response to said determining comprises
4 transmitting a flow control packet with zero credits
5 for each of a plurality of virtual lanes configured on
6 the port.

1 12. The method of claim 9, wherein said latency time
2 comprises an approximate amount of time required for
3 the port to transmit a maximum-sized Infiniband data
4 packet to the device.

1 13. The method of claim 9, wherein said latency time
2 comprises an approximate amount of time required for
3 the device to transmit a maximum-sized Infiniband data
4 packet to the port.

1 14. The method of claim 9, wherein said latency time
2 comprises an approximate amount of time required for
3 the device to respond to said transmitting flow control
4 credits to the device to stop transmission of the
5 packets in response to said determining.

1 15. The method of claim 8, wherein said size C is between
2 approximately one Kilobyte and approximately sixteen
3 Kilobytes.

1 16. The method of claim 1, further comprising:
2 buffering the packets transmitted by the device
3 subsequent to said determining in a reserved
4 amount of the portion of the memory, wherein said
5 reserved amount is beyond the predetermined
6 amount.

1 17. The method of claim 16, wherein said reserved amount is
2 between approximately eight Kilobytes and approximately
3 sixteen Kilobytes.

1 18. The method of claim 16, wherein said reserved amount is
2 based on an amount of data that may be transmitted to
3 the port during a latency time required to stop
4 transmission of the packets in response to said
5 determining.

1 19. The method of claim 18, wherein said latency time
2 comprises an approximate amount of time required for
3 the port to transmit a flow control packet for each of
4 a plurality of virtual lanes configured on the port.

1 20. The method of claim 18, wherein said latency time
2 comprises an approximate amount of time required for
3 the port to transmit a maximum-sized Infiniband data
4 packet to the device.

1 21. The method of claim 18, wherein said latency time
2 comprises an approximate amount of time required for
3 the device to transmit a maximum-sized Infiniband data
4 packet to the port.

1 22. The method of claim 18, wherein said latency time
2 comprises an approximate amount of time required for
3 the device to respond to said transmitting flow control
4 credits to the device to stop transmission of the
5 packets in response to said determining.

1 23. The method of claim 1, wherein said determining the
2 portion of the memory is filled a predetermined amount
3 comprises determining an amount of free space in the
4 portion of the memory drops below the predetermined
5 amount.

1 24. The method of claim 23, wherein said amount of free
2 space is between approximately eight Kilobytes and
3 approximately sixteen Kilobytes.

1 25. The method of claim 1, wherein said providing a portion
2 of a memory for buffering the packets comprises
3 dynamically allocating the memory from a pool of memory
4 shared among the port and a plurality of other
5 Infiniband ports.

1 26. The method of claim 1, wherein said providing a portion
2 of a memory for buffering the packets comprises
3 providing the memory in response to user input.

1 27. The method of claim 1, wherein said providing a portion
2 of a memory for buffering the packets comprises
3 providing the portion of the memory to the port based
4 on a plurality of other ports sharing the memory with
5 the port.

1 28. The method of claim 1, wherein said transmitting flow
2 control credits to advertise to the device buffering
3 resources of a size B comprises advertising at least
4 two maximum-sized Infiniband packets worth of flow
5 control credits for each of a plurality of virtual
6 lanes configured on the port.

1 29. The method of claim 1 further comprising:
2 configuring a plurality of virtual lanes on the port
3 prior to said transmitting flow control credits to

4 advertise to the device buffering resources of a
5 size B.

1 30. The method of claim 29, wherein a product of said
2 plurality of virtual lanes and a number of bytes
3 comprising two maximum-sized Infiniband packet exceeds
4 size A.

1 31. A method for controlling flow of packets into a
2 plurality of ports on an Infiniband device, comprising:
3 providing a memory for buffering the packets, wherein
4 the memory has a size A;
5 transmitting flow control credits by the plurality of
6 ports to advertise packet buffering resources of a
7 size B, wherein B is greater than A; and
8 transmitting flow control credits by at least one of
9 the plurality of ports to stop transmission of the
10 packets into the at least one port in response to
11 determining an amount of free space in the memory
12 drops below a predetermined threshold.

1 32. The method of claim 31, wherein said transmitting flow
2 control credits by the plurality of ports to advertise
3 packet buffering resources of a size B comprises
4 transmitting flow control credits for each of a
5 plurality of virtual lanes configured on each of the
6 plurality of ports.

1 33. The method of claim 31, wherein said predetermined
2 threshold is based on an amount of data that may be
3 transmitted to the plurality of ports during a latency
4 time required to stop transmission of the packets in
5 response to said determining.

1 34. The method of claim 31, wherein said predetermined
2 threshold is approximately zero, wherein said method
3 further comprises:
4 providing a second memory for buffering the packets
5 transmitted subsequent to said determining.

1 35. A system for buffering packets transmitted by a link
2 partner linked to an Infiniband port, comprising:
3 a first memory, for buffering the packets from the
4 port;
5 flow control logic, configured to advertise to the link
6 partner more buffering resources than are
7 available in said first memory for buffering the
8 packets if space is available in said first memory
9 to buffer the packets, and to advertise no
10 buffering resources if no space is available; and
11 a second memory, coupled between the port and said
12 first memory, for buffering the packets when no
13 buffering resources are available in said first
14 memory.

1 36. The system of claim 35, wherein said second memory is
2 configured to receive the packets independent of a
3 plurality of virtual lanes specified in the packets.

1 37. The system of claim 35, wherein a size of said second
2 memory is approximately an amount of data capable of
3 being transmitted to the port during a latency time
4 from when no buffering resources are available in said
5 first memory to when the link partner stops
6 transmitting the packets.

1 38. The system of claim 35, wherein said flow control logic
2 is configured to advertise to the link partner said
3 buffering resources for a plurality of virtual lanes
4 configured on the port.

1 39. A system for buffering packets transmitted by a link
2 partner linked to an Infiniband port, comprising:
3 a memory, having a size;
4 an inline buffer, coupled between the port and said
5 memory, for selectively buffering the packets if
6 said memory is full; and
7 flow control logic, configured to advertise to the link
8 partner more flow control credits than space
9 available in said memory, wherein said flow
10 control logic is further configured to advertise

11 to the link partner zero flow control credits when
12 said memory is full.

1 40. The system of claim 39, wherein said flow control logic
2 is configured to advertise to the link partner more
3 flow control credits than space available in said
4 memory across a plurality of virtual lanes configured
5 on the port.

1 41. A system for buffering packets transmitted by a link
2 partner linked to an Infiniband port, comprising:
3 a memory, for buffering the packets from the port;
4 a buffer controller, for monitoring an amount of free
5 space in said memory; and
6 flow control logic, configured to advertise to the link
7 partner more buffering resources than are
8 available in said memory for buffering the packets
9 from the port if said buffer controller indicates
10 said amount of free space is above a predetermined
11 threshold.

1 42. The system of claim 41, wherein said flow control logic
2 is further configured to advertise to the link partner
3 no buffering resources if said buffer controller
4 indicates said amount of free space is below said
5 predetermined threshold.

1 43. The system of claim 41, wherein said predetermined
2 threshold is approximately an amount of data capable of
3 being transmitted to the port during a latency time
4 from when said buffer controller indicates said amount
5 of free space is below said predetermined threshold to
6 when the link partner stops transmitting the packets.

1 44. The system of claim 41, wherein said flow control logic
2 is configured to advertise to the link partner said
3 buffering resources for a plurality of virtual lanes
4 configured on the port.

1 45. The system of claim 44, wherein said memory has a size,
2 wherein said plurality of virtual lanes configured on
3 the port multiplied by a size of at least two maximum-
4 sized Infiniband data packets substantially exceeds
5 said size of said memory.

1 46. An Infiniband device, comprising:
2 a plurality of ports, each having a plurality of
3 virtual lanes configured therein;
4 memory, for buffering packets received by said
5 plurality of ports, said memory having a
6 predetermined size; and
7 flow control, for advertising an amount of buffering
8 resources comprising at least two Infiniband

9 packets worth of flow control credits for each of
10 said plurality of virtual lanes configured in each
11 of said plurality of ports;
12 wherein said advertised amount of buffering resources
13 substantially exceeds said predetermined size of
14 said memory.

1 47. The device of claim 46, wherein said Infiniband device
2 is an Infiniband switch, router or channel adapter.

1 48. A buffering system in an Infiniband device, comprising:
2 a port, having a plurality of virtual lanes configured
3 therein;
4 a memory, for buffering packets received by said port,
5 said memory having a predetermined size; and
6 flow control, configured to advertise an amount of
7 buffering resources comprising at least two
8 Infiniband packets worth of flow control credits
9 for each of said plurality of virtual lanes
10 configured in said port;
11 wherein said advertised amount of buffering resources
12 substantially exceeds said predetermined size of
13 said memory.

1 49. The buffering system of claim 48, wherein said flow
2 control is further configured to advertise zero credits
3 for each of said plurality of virtual lanes configured

4 in said port upon determining less than a predetermined
5 amount of said memory is free to buffer said packets
6 received from said port.